Against Sequence Priming: Evidence from Constituents and Distituents in Corpus Data

Citation for published version:
Reitter, D & Keller, F 2007, Against Sequence Priming: Evidence from Constituents and Distituents in Corpus Data. in *Proceedings of the 29th Annual Conference of the Cognitive Science Society.* pp. 1421-1426. <http://csjarchive.cogsci.rpi.edu/proceedings/2007/docs/p1421.pdf>

Published In:
Proceedings of the 29th Annual Conference of the Cognitive Science Society

Against Sequence Priming: Evidence from Constituents and Distituents in Corpus Data

David Reitter (dreitter@inf.ed.ac.uk) and
Frank Keller (keller@inf.ed.ac.uk)
School of Informatics, 2 Buccleuch Place
Edinburgh EH8 9LW, UK

Abstract

Structural priming, i.e., the tendency to repeat linguistic material, can be explained by two alternative representational assumptions: either as the repetition of hierarchical representations generated by syntactic rules, or as the repetition of lexical sequences. We present two studies that test these explanations by investigating priming effects in a dialogue corpus. We compare syntactic constituents with distituents, i.e., part-of-speech pairs that cross constituent boundaries.

We find a reliable short-term priming effect for constituents, but no priming for distituents. This result supports the rule-based view of priming, which does not predict priming of distituents. The data are incompatible with a sequence priming analysis, which cannot distinguish between constituents and distituents. In a second corpus study, we study long-term priming and find priming effects for both constituents and distituents. This indicates that the mechanism underlying long-term adaptation is substantially different from short-term priming.

Keywords: language production, syntactic priming, implicit learning, corpus studies, constituents, distituents.

Introduction

When humans speak or write, they convert a conceptual representation of the message to be conveyed into sequences of sounds or letters. This task of *language production* is often analyzed in terms of a processing chain which includes conceptualization, formulation, and articulation (Levelt, 1989). The conceptualization module selects concepts to express, and the formulation module decides how to express them. Formulation involves determining the lexical, syntactic, and semantic representation of the utterance. Syntax determines the systematic relationship between meaning and form of an utterance, without which language could not be produced.

Given the central role of syntax in language production, it is not surprising that a significant amount of recent research has tried to establish the exact nature of the syntactic representations that underlie the production process. As syntactic structures cannot be observed directly, a number of indirect ways have been developed to investigate them. An important one is the study of *structural priming,* which is the preference of the language processor to re-use previous syntactic choices. As an example, consider the verb *give*, which can occur in either a double object (DO) construction (see (1-a)) or in a prepositional object (PO) construction (see (1-b)):

(1) a. The policeman gives a gun to the magician.
b. The policeman gives the magician a gun.

Experimental results (e.g., Bock 1986) show that participants who have a choice between producing the DO and the PO construction (e.g., in a picture naming task) are more likely to choose that construction which they (or their interlocutor) have produced previously.

Priming results such as this one give us a handle on syntactic representations: priming is only expected between constructions that share the same representation, therefore the presence or absence of priming can be used as a diagnostic for whether two construction involve identical representations or not. Using examples such as (1), it has been argued that priming takes place on the level of syntactic rules (though this can also be interpreted as priming of lexical sequences, as discussed below). There is also evidence for the priming of attachment decisions (Scheepers, 2003), and for the priming of sequences of constituents (Scheepers & Corley, 2000).

Recent corpus-based work has reinforced the structure-based view of priming. For example, Reitter et al. (2006a,b) demonstrated that priming can occur for arbitrary syntactic rules in a large collection of speech data. This is an important generalization of results from experimental work, which has only investigated priming for syntactic alternatives (such as (1) above), not for arbitrary rules.

However, the structure-based view of priming has been challenged by Chang et al. (2006), who propose a Simple Recurrent Network model that captures priming as the repetition of sequences of abstract lexical types, such as parts of speech. In this model, syntactic priming does not involve syntactic rules, but is explained simply as the learning of lexical sequences.

In this paper, we present corpus data that make it possible to directly compare the rule-based and the sequence-based view of priming. The key idea is to compare priming effects for *constituents* (i.e., linguistic units generated by syntactic rules) with priming effects for *distituents* (i.e., sequences of parts of speech that cannot form a linguistic unit). Only under the sequence-based account do we predict the priming of distituents.

Models of Syntactic Priming

Rule Priming

Traditionally, syntactic priming has been explained in terms of the activation of structural representations in the language production system (Bock, 1986; Branigan et al., 1999). In order to generate an utterance, a syntactic structure of this utterance has to be built, and this process involves the activation of syntactic frames, such as the double object frame of the verb *give* in (1-a). This activation decays over time, and when

Danielle S. McNamara and J. Greg Trafton, eds., *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, 1421–1426. Nashville, 2007.

A: And$_{\mathrm{CC}}$ all$_{\mathrm{DT}}$ of$_{\mathrm{IN}}$ a$_{\mathrm{DT}}$ sudden$_{\mathrm{JJ}}$ he$_{\mathrm{PRP}}$'s$_{\mathrm{HVS}}$ got$_{\mathrm{VBN}}$ a$_{\mathrm{DT}}$ hang$_{\mathrm{NN}}$ glider$_{\mathrm{NN}}$

B: I$_{\mathrm{PRP}}$ do$_{\mathrm{VBP}}$ n't$_{\mathrm{RB}}$ even$_{\mathrm{RB}}$ heard$_{\mathrm{VBN}}$ of$_{\mathrm{IN}}$ that$_{\mathrm{DT}}$ show$_{\mathrm{NN}}$

A: You$_{\mathrm{PRP}}$ have$_{\mathrm{VBP}}$ n't$_{\mathrm{RB}}$

B: It$_{\mathrm{PRP}}$'s$_{\mathrm{BES}}$ called$_{\mathrm{VBN}}$ McGyver$_{\mathrm{NNP}}$?

A: He$_{\mathrm{PRP}}$'s$_{\mathrm{BES}}$ like$_{\mathrm{UH}}$ a$_{\mathrm{DT}}$ semigovernment$_{\mathrm{JJ}}$ type$_{\mathrm{NN}}$ agent$_{\mathrm{NN}}$ who$_{\mathrm{WP}}$ the$_{\mathrm{DT}}$ Phoenix$_{\mathrm{NNP}}$ Foundation$_{\mathrm{NNP}}$ supposedly$_{\mathrm{RB}}$ ...

Figure 1: Excerpt from the tagged Switchboard data.

the production system has to generate another utterance, it is more likely to utilize a syntactic frame that has been pre-activated, i.e., that has been used in the recent past. This then leads to the priming effect, e.g., in the case of (1-a), the production system is more likely to generate another double object construction (rather than the alternative prepositional object construction in (1-b)).

The exact nature of the syntactic representations (syntactic frames, etc.) that underlie priming has been the subject of some debate. Recently, a number of corpus studies have provided evidence for *syntactic structure* as the correct level of representation. These demonstrated the repetition of syntactic rules in corpus data consistent with experimental results on syntactic priming. This includes evidence for the priming of specific constructions (Gries, 2005; Szmrecsanyi, 2005; Dubey et al., 2005) as well as evidence for a generalized priming effect that applies to arbitrary rules (Reitter et al., 2006a,b) and does not have to involve the alternation of semantically equivalent syntactic realizations (as in example (1)).

These corpus studies also constitute important corroborating evidence for the activation-based view, as they replicate the central characteristics of the experimental results on priming, including the rapid, exponential decay of the effect and the increased priming if head words are repeated (lexical boost) (Bock, 1986; Branigan et al., 1999).

Sequence Priming

Syntactic rules are not a necessary component of a model of sentence production, and recent modeling work has assumed that priming operates on sequences of abstract lexical categories rather than on rules. Many known priming effects can be explained in this way, e.g., the fact that (2-a) primes (2-c) could be due to the shared part of speech (POS) sequence NN DT PRP in both sentences. (See Table 1 for a subset of the part-of-speech categories used in this study, and Figure 1 for an excerpt from the corpus.) Sentence (2-b), on the other hand, contains a different POS sequence (NN DT NN) and therefore is expected to prime (2-d), but not (2-c), consistent with experimental results on the priming of prepositional object and double object constructions.

CC	coordinating conjunction (and, or)
DT	singular determiner/quantifier (this, that)
IN	preposition
JJ	adjective
MD	modal auxiliary (can, should, will)
NN	singular or mass noun
NNS	plural and/or possessive noun
PRP	personal pronoun
RB	adverb
UH	hesitation
VBZ	verb, 3rd. singular present
VBP	verb, present tense, other than 3rd singular
WDT	wh- determiner (what, which)
WP$	possessive wh- pronoun (whose)
WRB	wh- adverb (how, where, when)

Table 1: Common Brown/Switchboard part-of-speech tags.

Frequency	POS bigram	Frequency	POS bigram
38794	PRP VBP	38794	PRP VBP
12261	PRP VBD	25543	DT NN
6488	PRP MD	18275	IN DT
4443	NN PRP	15318	IN PRP
2901	WRB PRP	14623	NN IN
2868	PRP VBZ	12261	PRP VBD
2137	PRP VB	11561	JJ NN
1502	WDT PRP	10740	TO VB
1464	WP PRP	10017	CC PRP
1392	NNS PRP	9293	DT JJ
1162	VBD TO	8178	VBP PRP
1089	JJ TO	7838	RB RB
1031	PRP DT	7482	VBP RB
955	JJ PRP	7265	IN NN
827	PRP PRP	7137	RB JJ
604	VBN TO	6556	NN NN

Table 2: The most common *distituent* (left) and *constituent* (right) POS bigrams from the corpus.

(2) a. The$_{\mathrm{DT}}$ doctor$_{\mathrm{NN}}$ gives$_{\mathrm{VBZ}}$ some$_{\mathrm{DT}}$ flowers$_{\mathrm{NN}}$ to$_{\mathrm{IN}}$ his$_{\mathrm{PRP\$}}$ girl$_{\mathrm{NN}}$
b. The$_{\mathrm{DT}}$ doctor$_{\mathrm{NN}}$ gives$_{\mathrm{VBZ}}$ his$_{\mathrm{PRP\$}}$ girl$_{\mathrm{NN}}$ some$_{\mathrm{DT}}$ flowers$_{\mathrm{NN}}$
c. The$_{\mathrm{DT}}$ policeman$_{\mathrm{NN}}$ gives$_{\mathrm{VBZ}}$ a$_{\mathrm{DT}}$ gun$_{\mathrm{NN}}$ to$_{\mathrm{IN}}$ the$_{\mathrm{DT}}$ magician$_{\mathrm{NN}}$
d. The$_{\mathrm{DT}}$ policeman$_{\mathrm{NN}}$ gives$_{\mathrm{VBZ}}$ the$_{\mathrm{DT}}$ magician$_{\mathrm{NN}}$ a$_{\mathrm{DT}}$ gun$_{\mathrm{NN}}$

The sequencing view of priming is central to Chang et al.'s (2006) Dual-path Model, a connectionist model of sentence production that aims to account for results from both language acquisition and syntactic priming. At the core of the Dual-path Model there are two mechanisms. The first one is the Sequencing System, consisting of a Simple Recurrent Network (SRN, Elman 1990) which generates sequences of words, compressing them to abstract parts of speech cat-

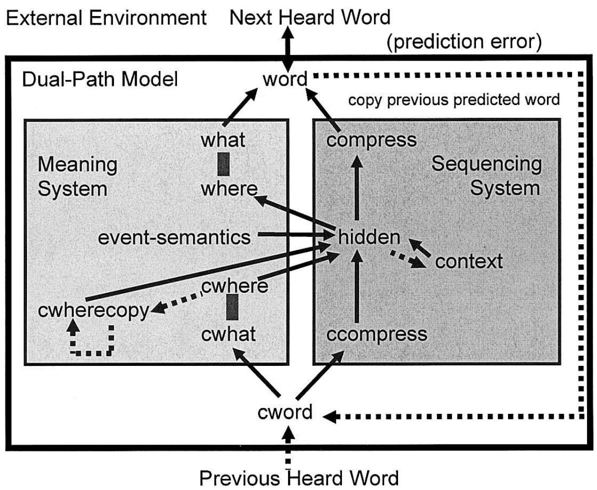

Figure 2: Schematic view of the Dual-path Model (figure from Chang et al. 2006)

egories. As is common for SRNs, language production is viewed as the task of predicting the next word given its left context, and an error-driven learning algorithm is used to train the model. The second mechanism in the Dual-path Model is the Meaning System which maps meaning representations to words and vice versa. These representations consist of *what*- and *cwhat*-nodes (representing the lexical semantics of words in production and comprehension, respectively) and *where*- and *cwhere*-nodes (representing words' semantic roles in production and comprehension). Figure 2 gives a schematic view of the Dual-path Model. Note that the model contains a self-monitoring loop which connects the currently produced word with the comprehended version of the previously produced word (*cword* in the diagram).

The Dual-path Model accounts for a wide range of structural priming results, as well as for certain findings in the language acquisition literature (preferential looking and elicited production studies). The model makes two key assumptions, which we will test in the present paper: (1) language comprehension and production is based on abstract lexical (POS) sequences as the underlying representation, no hierarchical syntactic structures (and no syntactic rules) are involved; (2) the same implicit learning processes underlie language production and acquisition, which means that short-term priming (which decays in a matter of seconds) and long-term priming (which can take days to decay) are accounted for by the same mechanism, i.e., an SRN trained to predict POS sequences.

Distituents

In order to distinguish structural from sequential priming, we use the notion of *distituents* from the grammar induction literature (e.g., Magerman & Marcus 1990; Kuhn 2004). Distituents are pairs of POS tags that cannot form a syntactic unit. All other pairs, i.e., the ones that occur in a syntactic unit, are deemed *constituents*. Crucially, such constituent pairs are predicted to show decaying repetition due to priming under both assumptions, structural and sequential priming. *Distituent* pairs, however, will show priming only if sentence production is based on sequential representations. Thus, under the structure-based view, there should be no distituent priming, as distituents (by definition) cannot be generated by syntactic rules.

To define distituents more precisely, we refer to the POS categories and the tree-structured syntactic analysis of each sentence. The syntax tree then defines *constituents* or subtrees. For example, in the syntax tree in (3), *the policeman*, among other phrases, forms a constituent.

(3) [[The$_{DT}$ policeman$_{NN}$] [shows$_{VBZ}$
[$_i$ the$_{DT}$ girl$_{NN}$] [$_j$ his$_{PRP\$}$ gun$_{NN}$]]]

A distituent is a POS pair that cannot be adjacent without crossing at least one constituent boundary. For example, NN PRP$ (noun, possessive pronoun) is a distituent in English, because there can be no constituent that directly combines a noun followed by a possessive pronoun. Of course, such a POS sequence still occurs in the data (as in (3)), but for a distituent bigram, the two POS tags will always belong to two *different* constituents (in the above case the two argument noun phrases *i*, *j*). To give another example, DT NN is not a distituent, because the determiner and the noun directly form a noun phrase. NN VBZ is not a distituent either: while it does cross constituent boundaries in (3), it appears without doing so (in its own constituent) in a verbal phrase with an intransitive verb elsewhere in the corpus (e.g., [*before* [*school starts*]]). Table 2 lists the most frequent distituents.

An equivalent definition of distituency refers to dominance in the syntax tree: Two adjacent POS tags α, β are distituent if and only if there is no node *N* such that *N* immediately dominates α and β, and all other instances of α, β in the corpus are distituents.[1]

If a corpus of syntactically annotated material is available, then the syntactic annotation can be used to identify distituents in the data as follows: for every sequence of two adjacent parts of speech (bigram) in the corpus, we determine whether it occurs inside a constituent without crossing constituent boundaries anywhere in the corpus. If this never happens in the corpus, then we regard this sequence as a distituent. Note that distituents (contrary to constituents) do not have a hierarchical structure – they should be regarded simply as bigrams that cross constituent boundaries.

Experiment 1: Short-term Priming

If Chang et al.'s (2006) sequencing view of priming is correct, then there should be no systematic difference between constituents and distituents. Therefore, his model predicts that in corpus data, we should find priming for both constituents and

[1] It should be noted that this definition invokes immediate dominance, i.e., it leaves open the possibility that the distituent is part of a larger constituent that dominates it, but not immediately. Strictly speaking, under a rule based view we would therefore expect less priming for distituents, rather than no priming at all.

for distituents. On the other hand, if the rule-based view is correct, then priming should be confined to constituents, as distituents cannot be generated by syntactic rules, and therefore cannot be subject to priming. The present experiment tests these two alternative hypothesis for short-term priming, i.e., for structural repetition that decays rapidly.

Method

Data The Switchboard corpus is a large data set of spontaneous conversations between over 500 participants, speaking varieties of North American English. We use a subset of 426 conversations averaging 6 minutes in length, which have been transcribed and then syntactically annotated with syntax trees. Exact timing information is available is each word (and therefore for each constituent).

Distituents were identified in the Switchboard corpus following the definition given in the previous section. Bigrams including hesitations such as *like* and *uh*, or with POS tags not identified by the original annotation (marked XX), were excluded. This way, we extracted 378 different types of POS bigrams, 80 of which were distituents. (See Table 2 for common distituents and constituents.) Data points with rare POS sequences ($f \leq 10$) and unknown POS tags were discarded.

Statistical Analysis To analyze priming effects in our corpus data, we examine the repetition of POS sequences. Whenever a POS sequence is repeated within a short time period more often than we would expect from chance repetition, we accept it as an example of structural priming.

As discussed before, short-term priming is subject to a swift decay. The increase in repetition probability is seen shortly after the stimulus, but less so a few seconds later. Therefore, we use the time elapsed after a stimulus to predict whether repetition will occur. A logistic regression model was used to compute a correlation coefficient between repetition and the temporal distance d (as co-variate TIME).

For each occurrence of a POS sequence (*target*) at a time t, we examine the POS sequences in the one-second time period $[t-d-0.5,\ t-d+0.5]$. If the parts of speech re-occur, we count the target occurrence as *primed*, otherwise as control case. This is the predicted binary variable, PRIMED.

If there is no structural priming effect (null hypothesis), we would expect there to be no relationship between PRIMED and TIME. An interaction between this effect and the factor distinguishing distituent from constituent bigrams (DISTITUENT) would reveal differences in priming strength between constituents and distituents.

To account for frequency effects in priming as they have been reported previously, we include the normalized bigram frequency as a co-variate POSFREQ. A further factor TYPE distinguishes priming between speakers (comprehension-production priming, CP) from priming within a speaker (production-production priming, PP): only in the latter case were prime and target uttered by the same speaker.

To implement this logistic regression model, we use generalized linear mixed models with a logit-link, conservatively grouping sequences stemming from the same utterance to reflect potential inter-dependence due to syntactic constraints. The data set was randomly balanced with respect to the response variable in the respective experiment.[2]

Interactions (and main effects) were removed where appropriate, i.e., where there was no significant coefficient and no dependent interaction.

Results

The results show a reliable main effect for log(TIME) (decay $\beta = -0.067$, $p < 0.0001$), indicating a baseline priming effect. The model also showed a reliable interaction of log(TIME) and DISTITUENT ($\beta = 0.183$, $p < 0.05$), indicating reliably less priming for distituents. The fact that the sum of the two coefficients is positive, indicates that there is no decay: $-0.067 + 0.183 > 0$, which means that there is in fact no priming for distituents.

log(TIME) also interacts reliably with log(POSFREQ) ($\beta = 0.156$, $p < 0.0001$), showing that higher-frequency POS bigrams receive less priming. We see a small but reliable interaction of log(TIME) and TYPE ($\beta = 0.050$, $p < 0.001$), indicating that priming is weaker between speakers than within.[3]

Discussion

The main priming effect we found is consistent with the experimental literature (Bock, 1986; Pickering & Garrod, 2004). We also replicated the priming, frequency, and type effects found in previous corpus studies on syntactic priming (Reitter et al., 2006b) and frequency effects found experimentally for relative clause attachment priming (Scheepers, 2003).

With respect to the hypothesis leading to this experiment, we found not only reliably less priming for distituents: the decay coefficient for distituents was numerically estimated to be positive, i.e., we see no priming for these part of speech sequences. This provides evidence against a non-structural priming account.

Given the marked contrast between constituents and distituents, we can argue for an architecture of the human formulation mechanism that involves hierarchical syntactic representations. Obviously, this does not exclude the possibility of memory effects involving surface-structure word or POS sequences. The next experiment examines this question.

Experiment 2: Long-term Priming

Classical priming effects are strong (up to 40 percent above the baseline for passives; around 10 percent for syntactic rules, Reitter et al. 2006b). They decay quickly (Branigan et al., 1999) and reach a low plateau after a few seconds.

[2] In an experimental design, we would control and balance dependent variables rather than the response, but here, where we analyze interested in the fitted interactions, the model fitting is more reliable with a balanced data set.

[3] Further effects, irrelevant to the experiment because they model chance repetition as well as repetition when primed, were: log(POSFREQ) ($\beta = 0.45$, $p < 0.0001$), TYPE=CP ($\beta = -0.19$, $p < 0.0001$), and DISTITUENT ($\beta = -0.81$, $p < 0.0001$).

Such syntactic short-term priming is similar to what has been shown in lexical priming studies (e.g., Swinney et al. 1979). What complicates matters is that there is also a longer-term repetition effect that has been reported in the literature.

Adaptation, also termed long-term priming, has been shown to last longer, from minutes (Bock & Griffin, 2000) to several days. Lexical boost effects, in which the lexical repetition strengthens structural priming, have been observed for short-term priming, but not for long-term priming trials where material intervenes between prime and target utterances (Konopka & Bock, 2005). Thus, short-term and long-term adaptation effects may be due to separate cognitive processes, as recently argued by Ferreira & Bock (2006). Short-term priming is arguably a mechanistic effect related to language processing, while adaptation is more similar to a implicit learning in that it lacks strong decay. If priming and adaptation are indeed two qualitatively different cognitive processes, then Chang's Dual-path Model may be able to account for adaptation. This would require that learning applies to sequences rather than structures. Therefore, comparing the adaptation of constituent and distituent bigrams would shed light on this question. This is the aim of the present experiment.

Method

The data set was the same as in Experiment 1.

While short-term priming can be pin-pointed using the characteristic decay, for long-term priming we need to inspect whole dialogues. As in Experiment 1, we use a binary response variable PRIMED to reflect the repetition of a POS sequence. While we estimated PRIMED as a function of distance between prime and target in Experiment 1, with primes occurring in a one-second priming period at a set distance before the target, we now regard the first half of a dialogue as priming period, testing all POS sequences in the second half for repetition.

We will contrast PRIMED in two conditions, which distinguish situations where priming can take place (SAMEDOC=1) from others, where repetition is only due to chance (SAMEDOC=0). To do so, we split each dialogue into two equal halves, but exclude a 10-second portion in the middle to avoid short-term priming effects. The first half is designated as priming half, the second half contains the targets. For each target POS bigram, we check whether it has already occurred in the priming half (PRIMED=1).

For the priming condition SAMEDOC=1, we keep dialogues together: priming and target halves stem from the same original dialogue. For the non-priming control condition (SAMEDOC=0), priming and target halves are randomly chosen so that they stem from different dialogues.

We can then cast long-term adaptation as the differential between rule repetition in document halves of single dialogues, and repetition in dialogues halves sampled from different dialogues. The goal is now to establish a main effect of SAMEDOC for adaptation, and its interaction with DISTITUENT.

Results

The resulting model shows a number of reliable main effects and interactions. In the following, we will not only analyze significance, but also pay attention to effect sizes.

We find a reliable main effect of SAMEDOC ($\beta = -0.34$, $p < 0.0001$) and an interaction of log(POSFREQ) with SAMEDOC ($\beta = -0.15$, $p < 0.0001$). This indicates that at low bigram frequencies (log(POSFREQ) < -2.27), repetition of constituents is greater in priming dialogues than in the control. We find positive adaptation of constituent bigrams.

Further, the model shows reliable interaction of DISTITUENT with SAMEDOC ($\beta = -0.38$, $p < 0.05$) and with SAMEDOC:log(POSFREQ) (triple interaction). This means that at similarly low bigram frequencies (log(POSFREQ) < -2.56), again, repetition of distituents is greater in priming dialogues than in the control. We thus find positive adaptation of distituent bigrams.

Centered and transformed bigram frequencies range from -6.67 to 1.50, with mean $\mu(\log(\text{POSFREQ})) = -0.81$, standard deviation $\sigma(\log(\text{POSFREQ})) = 1.48$, and the lower quartile at -1.7160. The above adaptation effects apply to the 13% of bigrams with the lowest frequencies.[4]

The model shows positive adaptation for low-frequency bigrams, both in the cases of constituents and distituents. This evidence is supported further by a simplified model, where the triple interaction involving the POS frequency is removed. In this simplified model, there is no reliable interaction effect of DISTITUENT and SAMEDOC can be found ($p = 0.38$).

We conclude that there is no evidence for a difference in long-term adaptivity between constituents and distituents.

Discussion

Short-term priming, decaying within a few seconds, and long-term adaptation lasting minutes and in some cases even days, differ substantially (see Ferreira & Bock 2006). Our data show both kinds of repetition effects. However, syntactic structure clearly mattered only for short-term processing effects: long-term adaptation appears to operate on abstract lexical sequences rather than syntactic structure.

A model where sequences of part-of-speech or lexemes are memorized as procedures would explain the findings. Effectively, this likens long-term adaptation to a procedural memory effect. Stored procedures can certainly help speakers to produce and listeners to understand language, and they may support alignment effects in dialogue (Pickering & Garrod, 2004), and they are in line with Chang et al.'s (2006) model. So while we argue against the sequential account for priming, we believe it to be plausible for long-term adaptation processes.

[4] Further coefficients were fitted which are irrelevant to our purposes because they describe effects on chance repetition: log(POSFREQ) ($\beta = 1.73$, $p < 0.0001$), DISTITUENT ($\beta = -1.02$, $p < 0.0001$), log(POSFREQ):DISTITUENT ($\beta = -0.45$, $p < 0.0001$).

Conclusions

The aim of this paper was to shed light on the representations that underlie the human language production system by investigating the well-know structural priming effect that occurs when humans generate sentences. Structural priming, i.e., the repetition of previously used linguistic structures, can be explained using at least two alternative representational assumptions: either as the repetition of hierarchical representations generated by syntactic rules as proposed by Bock (1986) and Branigan et al. (1999), or as the repetition of sequences of abstract lexical representations (e.g., parts of speech) as proposed by Chang et al. (2006).

We presented data from two studies designed to distinguish the rule-based view from the sequencing view for priming. We investigated priming effects in a dialogue corpus for two types of part-of-speech pairs: Constituent POS pairs, which can occur within a syntactic constituent generated by a syntactic rule, and distituent POS pairs, which cross constituent boundaries and can never occur within a constituent.

Experiment 1 dealt with short-term priming, i.e., with repetition effects that decay within a few seconds. We found a reliable priming effect for constituents bigrams, but not for distituent bigrams. This finding is compatible with the structure-based view of priming, which would not expect priming of distituents, as these cannot be generated by syntactic rules. The results are at odds with the sequence priming view, which cannot distinguish between constituents and distituents, and would therefore predict priming for both.

Experiment 2 extended the study of syntactic priming to long-term adaptation effects. This repetition bias remains over long periods of time (hours and days) and its characteristics differ from those of short-term priming (e.g., no lexical boost). Our corpus study found a reliable long-term adaptation effect for low-frequency bigrams, which was similarly strong for distituents. This implies that the mechanisms underlying long-term adaptation and short-term priming differ.

Overall, our results are difficult to accommodate by simulations of sentence production such as the Dual-path Model, which assumes sequence-based sentence production and does not involve a notion of constituency, and therefore cannot explain the lack of short-term priming for distituents. Also, the Chang et al. (2006) model assumes a generalized implicit learning mechanism underlies both short-term and long-term priming. Again, this is at variance with our findings, which show clear difference between the two effects. Finally, we note that there are also experimental results, such as the priming of relative clause attachments (Scheepers, 2003) that are puzzling for the sequence-based view, as both high and low attachment involve the same POS sequence.

We conclude that an empirically adequate model of syntactic priming has to invoke a mechanism that operates on hierarchical syntactic representations to explain short-term priming, while a separate mechanism (perhaps implicit sequence learning) has to be invoked to explain long-term priming. This is consistent with a rule-based view of priming. Priming operates on a time span in which syntactic analysis in comprehension and syntactic realization in language production are affected. Adaptation is a memory effect, and simple sequences of linguistic representations may be implicitly learned.

References

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387.

Bock, J. K., & Griffin, Z. (2000). The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*, 177–192.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in language production: Evidence for rapid decay. *Psychonomic Bulletin and Review*, *6*, 635–640.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*, 234–272.

Dubey, A., Sturt, P., & Keller, F. (2005). Parallelism in coordination as an instance of syntactic priming: Evidence from corpus-based modeling. In *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing*, (pp. 827–834), Vancouver.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Ferreira, V., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*.

Gries, S. T. (2005). Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, *35*.

Konopka, B., & Bock, J. K. (2005). Helping syntax out: What do words do? In *Abstract at 18th CUNY*, Tucson, Arizona.

Kuhn, J. (2004). Experiments in parallel-text based grammar induction. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*, (pp. 470–477), Barcelona.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Magerman, D. M., & Marcus, M. P. (1990). Parsing a natural language using mutual information statistics. In *Proceedings of the 8th National Conference on Artificial Intelligence*, (pp. 984–989), Cambridge, MA. AAAI Press.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–190.

Reitter, D., Keller, F., & Moore, J. (2006a). Computational modelling of structural priming in dialogue. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers*, (pp. 121–124), New York.

Reitter, D., Moore, J., & Keller, F. (2006b). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, (pp. 685–690), Vancouver.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition*, *89*, 179–205.

Scheepers, C., & Corley, M. (2000). Syntactic priming in German sentence production. In L. R. Gleitman, & A. K. Joshi (eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, (pp. 435–440), Mahwah, NJ. Lawrence Erlbaum Associates.

Swinney, D., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across modalities in the processing of individual words and sentences. *Memory and Cognition*, *7*, 159–165.

Szmrecsanyi, B. (2005). Creatures of habit: A corpus-linguistic analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory*, *1*, 113–149.